

Large Scale Evaluation of Relationships between Hydrologic Signatures and Processes

Hilary K McMillan¹, Sebastian J. Gmann², and Ryoko Araki¹

¹San Diego State University

²University of Potsdam

November 24, 2022

Abstract

Dominant processes in a watershed are those that most strongly control hydrologic function and response. Estimating dominant processes enables hydrologists to design physically realistic streamflow generation models, design management interventions, and understand how climate and landscape features control hydrologic function. A recent approach to estimating dominant processes is through their link to hydrologic signatures, which are metrics that characterize the streamflow timeseries. Previous authors have used results from experimental watersheds to link signature values to underlying processes, but these links have not been tested on large scales. This paper fills that gap by testing signatures in large sample datasets from the U.S., Great Britain, Australia, and Brazil, and in Critical Zone Observatory (CZO) watersheds. We found that most inter-signature correlations are consistent with process interpretations, i.e., signatures that are supposed to represent the same process are correlated, and most signature values are consistent with process knowledge in CZO watersheds. Some exceptions occurred, such as infiltration and saturation excess processes that were often misidentified by signatures. Signature distributions vary by country, emphasizing the importance of regional context in understanding signature-process links and in classifying signature values as ‘high’ or ‘low’. Not all signatures were easily transferable from small- to large-scale studies, showing that visual or process-based assessment of signatures is important before large-scale use. We provide a summary table with information on the reliability of each signature for process identification. Overall, our results provide a reference for future studies that seek to use signatures to identify hydrological processes.

Large Scale Evaluation of Relationships between Hydrologic Signatures and Processes

H. McMillan^{1*}, S. J. Gnann², and R. Araki¹

¹Department of Geography, San Diego State University, San Diego, CA, USA

²Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany

*Corresponding author: Hilary McMillan (hmcmillan@sdsu.edu)

Key Points:

- Large sample signature distributions enabled us to put signature values into context as high or low, but differ by country.
- Most signatures agreed with the processes they are supposed to represent, except for infiltration and saturation excess signatures.
- We provide a table of signatures with recommendations on their reliability and use for process interpretation.

Abstract

Dominant processes in a watershed are those that most strongly control hydrologic function and response. Estimating dominant processes enables hydrologists to design physically realistic streamflow generation models, design management interventions, and understand how climate and landscape features control hydrologic function. A recent approach to estimating dominant processes is through their link to hydrologic signatures, which are metrics that characterize the streamflow timeseries. Previous authors have used results from experimental watersheds to link signature values to underlying processes, but these links have not been tested on large scales. This paper fills that gap by testing signatures in large sample datasets from the U.S., Great Britain, Australia, and Brazil, and in Critical Zone Observatory (CZO) watersheds. We found that most inter-signature correlations are consistent with process interpretations, i.e., signatures that are supposed to represent the same process are correlated, and most signature values are consistent with process knowledge in CZO watersheds. Some exceptions occurred, such as infiltration and saturation excess processes that were often misidentified by signatures. Signature distributions vary by country, emphasizing the importance of regional context in understanding signature-process links and in classifying signature values as ‘high’ or ‘low’. Not all signatures were easily transferable from small- to large-scale studies, showing that visual or process-based assessment of signatures is important before large-scale use. We provide a summary table with information on the reliability of each signature for process identification. Overall, our results provide a reference for future studies that seek to use signatures to identify hydrological processes.

1 Introduction

1.1 Hydrologic function and dominant processes

Within the hydrologic cycle, watersheds transport water from the land surface to its release as river flow, evapotranspiration, or groundwater. This role is referred to as “watershed function” and can be divided into key categories, such as partitioning, storage, and release of water (Black, 1997; McDonnell & Woods, 2004; Wagener et al., 2007). For example, partitioning includes interception, infiltration, percolation, runoff, and return flow processes. Storage includes snow, unsaturated or saturated zone storage, perched or deeper aquifers, and lakes. Release of water includes evapotranspiration, channel flow, and groundwater flow out of the watershed. Inherent in watershed process descriptions is the idea of “dominant processes.” Although watersheds might include a wide variety of processes under certain conditions, dominant processes are those most influential in controlling the hydrologic function and response (Grayson & Blöschl, 2001). For example, infiltration and saturation excess processes may both occur in a watershed, but the dominant process is the one that most strongly controls the magnitude and shape of the hydrograph.

There are many reasons to estimate the dominant processes in a watershed. Identifying the processes is a first step to developing models that provide physically realistic simulations (Grayson & Blöschl, 2001; Gupta et al., 2014). This is important given a new generation of hydrologic models with flexible structures that can simulate spatially variable processes, but may lack the corresponding spatial process knowledge (Clark et al., 2015). Watershed managers can apply process knowledge when designing interventions to intercept floodwater or prevent polluted runoff. More fundamentally, hydrologists seek to explain how climate, landscape, and critical zone

features control watershed processes and runoff generation (Dunne, 1978; Fan et al., 2019; Sivapalan, 2006). To achieve these goals, accurate estimation of dominant processes is essential.

1.2 Estimating dominant processes using regionalization and modeling

Two main approaches have been used to estimate dominant processes: regionalization and modeling. In the regionalization approach, knowledge is used from other similar watersheds. For example, Peschke et al. (1999) propose a regionalization method based on their experience in two experimental watersheds. After a literature review of conditions that favor the process of interest, they examined hydrograph shapes in the target basin and compared these with nearby basins. They examined potential runoff contributions from different land covers based on water balance estimates. The approach culminated in a rules-based assessment of which processes are possible given hydrologic and landscape characteristics. This method was automated and applied in a mesoscale basin by Hellie et al. (2002) to create process-oriented subdivisions. A similar, decision-tree approach was created by Scherrer & Naef (2003) to identify processes in highly-instrumented plots, providing a structured method to translate hydrologic observations into dominant process identification.

In the modeling approach, dominant processes are those which show the greatest sensitivity and improvement when incorporated into a model. Each candidate process can be added into the model in turn, and the model tested for improved performance (Sivakumar, 2008). At the same time, the dimensionality of the system can be analyzed to estimate how many processes are needed, although this approach cannot identify specific processes. Alternatively, a model which already incorporates all the candidate processes can be used in a sensitivity analysis. Markstrom et al. (2016) undertook a U.S.-wide assessment of parameter sensitivity for the PRMS model. Model parameters were grouped by process, and for each hydrological response unit, dominant processes were those with the highest sensitivity scores in their related parameters. This method was used to produce US-wide maps of process importance.

1.3 Hydrologic Signatures link to processes

A new approach to estimating dominant processes is through their link to hydrologic signatures. Signatures are quantitative metrics that describe statistical or dynamical features of streamflow timeseries, and are often used to assess model ability to simulate streamflow dynamics (McMillan, 2021). Examples include annual flood, baseflow index, slope of the flow duration curve, and descriptors of recession shapes. Some hydrologic signatures have well-understood links to processes in the upstream watershed, such as hydrograph recession shapes that can be derived from watershed storage-discharge behavior.

Such signature-process links have been used to assess dominant processes. Beighley et al. (2005) made a qualitative assessment of dominant processes based on total runoff ratio, hydrograph recession rate and flashiness, and change in event runoff ratio with season. They checked that the inferred processes were plausible given soil depth and impervious area, and added spatial data to understand process distribution within the watershed. The processes were then included in a watershed model. Recently, Wu et al. (2021) inferred patterns of runoff generation processes in the U.S. by using six signatures to cluster watersheds into eight classes. Signature values in each class were used to identify the dominant processes as infiltration excess, saturation excess/subsurface stormflow, lateral preferential flow, or baseflow. The classes were linked to physical watershed characteristics using random forest modeling. Both these studies rely on

proposing links between signatures and processes, which then enable a regionalization-type approach to estimate dominant processes.

A catalog of signature-process links was created by McMillan (2020) who collected streamflow signatures and independent information on dominant processes in papers from experimental watersheds. However, these links might be specific to particular climates or hydrologic regimes. To provide a sound basis for large-scale estimation of processes based on streamflow data, we must be sure that these links are consistent across watersheds. Some evidence for consistency exists, e.g., storage-related signatures (baseflow index and watershed sensitivity to runoff) were consistently linked to regolith development (weathering and creation of clay lenses) across the U.S. Critical Zone Observatory network (Wlostowski et al., 2020). However, further evidence is required to establish consistency of interpretation for a wide range of signatures and watershed characteristics.

1.4 Aims of the paper

The aim of this paper is to determine whether links between streamflow generation processes and streamflow signatures are consistent across a large sample of watersheds, or to determine for which signatures and processes the links hold. Understanding if and where the signature-process relationship is consistent will enable us to choose robust and reliable signatures to estimate dominant processes from large databases of streamflow data.

2 Data

We used two sources of hydrologic data: large scale data from several CAMELS datasets and small scale data from several Critical Zone Observatory (CZO) watersheds. Watershed locations are illustrated in Figure 1.

2.1 CAMELS datasets

We used several CAMELS datasets to test whether large-scale correlations and patterns in signature values conform with process knowledge. CAMELS datasets are national or continental-scale datasets of daily streamflow and forcing climate variables for watersheds, mostly with low influence from human impacts. We used CAMELS data from the U.S. (Addor et al., 2017; Newman et al., 2015), Great Britain (Coxon et al., 2020), Australia (Fowler et al., 2021) and Brazil (Chagas et al., 2020), and hourly CAMELS U.S. rainfall data from Gauch et al. (2020; 2021).

We used CAMELS streamflow (Q), precipitation (P), and potential evapotranspiration (PET) data for water years 1989 to 2009 and only kept watersheds with at least 99% complete records. Water years are defined as starting from 1 October for the U.S. and Great Britain, 1 April for Australia, and 1 September for Brazil. A few watersheds had very small negative PET values, and those were set to zero. We removed watersheds with more than 30% of precipitation falling as snow, which is about the upper limit of the CZO data investigated here. We also removed watersheds with significant human influences based on the following criteria, noting that this had a very small impact on the results. The number of watersheds used is shown in brackets:

- CAMELS U.S.: we kept all watersheds as they are near-natural (546 watersheds),
- CAMELS Great Britain: we only used benchmark watersheds (120 watersheds),

- CAMELS Australia: we removed watersheds with a river disturbance index > 0.2 (87 watersheds),
- CAMELS Brazil: we removed watersheds with `consumptive_use_perc` $> 5\%$ and watersheds with `regulation_degree` $> 10\%$ (486 watersheds).

2.2 CZO datasets

Critical zone observatories are highly-instrumented watersheds that are used to study interconnected hydrological, physical, biological, and chemical processes at the Earth's surface. These observatories offer precipitation, climate and streamflow data, and extensive literature describing hydrological processes, which can be compared with processes inferred from signature values. We conducted a signature analysis at five CZO sites with a total of eight streamflow gauges: Eel River, California (Elder and Dry Creeks); Shale Hills, Pennsylvania (Shale Hills Creek); Luquillo, Puerto Rico (Rio Mameyes and Rio Icaros); Intensively Managed Landscapes (IML), Illinois/Iowa (Upper Sangamon River); Santa Catalina, Arizona (Marshall Gulch and Oracle Ridge streams).

These observatories encompass a wide range of hydrological and climatological conditions, from arid, mountainous landscapes in Arizona, to tropical forest in Puerto Rico, to a humid, steep watershed in Pennsylvania. The CZOs with paired sites offer the opportunity to compare signature values in contrasting sites under similar climate conditions. In particular, Elder and Dry Creeks, and Rios Mameyes and Icaros differ significantly in underlying geology. These five observatories were selected from the CZO network as those with less than 30% of precipitation falling as snow, because event-based signatures in particular are unreliable under high snowfall conditions. Future options for including landscapes with significant snow could be to exclude snowmelt periods from the analysis (although current methods for determining spring snowmelt onset do not perform well in rivers with winter rains (Lundquist et al., 2004)); or to run a snow model to simulate soil water input, although this might lead to unwanted signature dependence on model characteristics.

For each site, raw data were processed into precipitation, streamflow, and PET time series with consistent hourly and daily timesteps, with the exception of Luquillo for which only daily data was available. Data were obtained from the Level 1 streamflow, precipitation, and meteorological datasets provided by Wlostowski et al., (2020), which comprise re-formatted versions of raw data. For Luquillo, we used Level 2 precipitation datasets that had been corrected based on annual totals (Wlostowski et al., 2020). For Eel River, additional streamflow and precipitation data for the neighboring Dry Creek were provided by D. Dralle (*pers. comm.*). Where multiple precipitation gauges were available, we calculated areal averages following the site-specific methods described by Wlostowski et al. (2020). Where necessary, we used disaggregated daily streamflow values from United States Geological Survey (USGS) gauges to infill missing hourly data. PET values were calculated from meteorological variables using the algorithm described by Zotarelli et al. (2010).

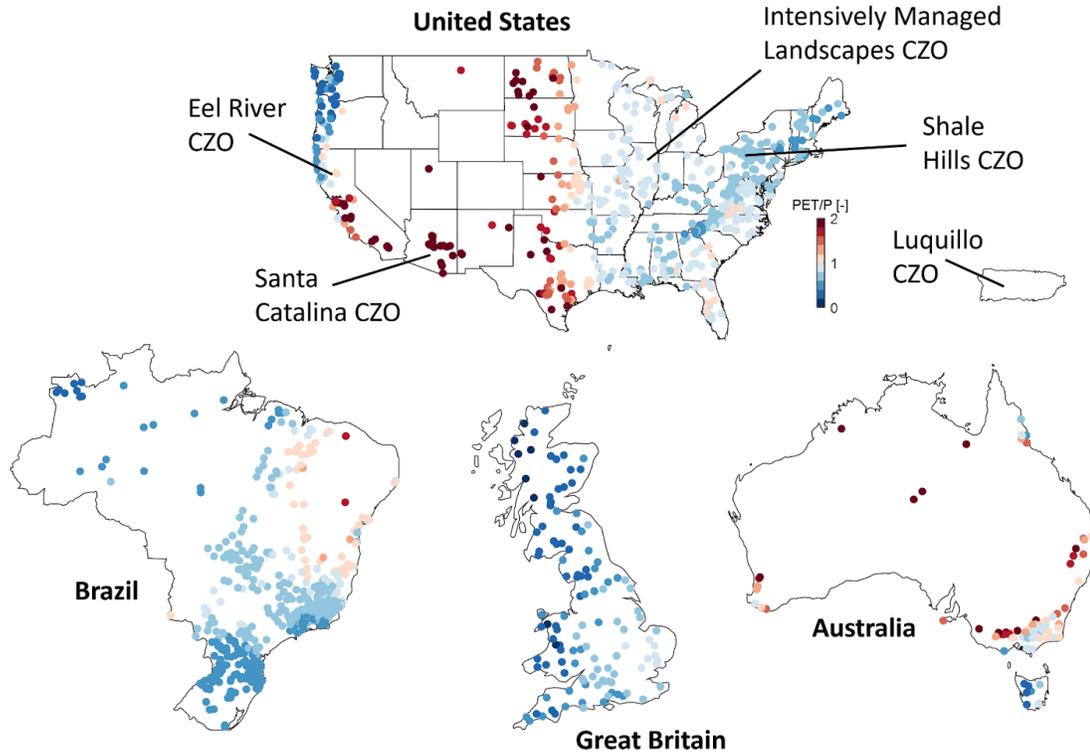


Figure 1. Map of CAMELS watersheds colored according to their aridity index (PET/P) and locations of CZOs. Note that the maps of the countries are not to the same scale.

3 Hydrologic signatures

The signatures tested in this paper relate to baseflow/groundwater processes and overland flow (saturation and infiltration excess) and are taken from the McMillan (2020) catalog. Matlab codes to calculate these signatures were implemented as part of the Toolbox for Streamflow Signatures in Hydrology (TOSSH) (Gnann et al., 2021a). The TOSSH toolbox provides standardized methods for hydrologic signature calculations, including recommended parameter values. Minor changes from the original catalog were made to revise or remove three signatures for ease of interpretation (see Table S1 in the supporting information). Full descriptions of the signatures calculated for this paper are given in Table 1 (17 signatures for groundwater/baseflow processes and 9 signatures for overland flow processes), and their Matlab code can be found at https://tosstoolbox.github.io/TOSSH/p2_signatures.html#process-based-signature-sets.

Table 1. Groundwater and overland flow signatures used in this paper.

Signature	Unit	Description
Groundwater and baseflow signatures		
TotalRR	-	Total runoff ratio
EventRR	-	Event runoff ratio (average over all events)
RR_Seasonality	-	Runoff ratio seasonality (summer total RR/winter total RR)
StorageFraction	-	Ratio between active and total storage
ActiveStorage	mm	Active storage defined as maximum storage deficit

TotalStorage	mm	Total storage calculated by extrapolation to find storage deficit at near-zero flow
Recession_a_Seasonality	-	Seasonal variations in recession 'a' parameter, related to recession timescale.
AverageStorage	mm	Average storage derived from average baseflow and storage-discharge relationship
RecessionParameters_b	-	Recession analysis parameters (T0, b) approximate storage-discharge relationship. b is a shape parameter.
RecessionParameters_T0	d	Characteristic timescale of recessions, at median flow.
MRC_num_segments	-	Number of different segments in nonparametric master recession curve (MRC)
BFI	-	Baseflow index, i.e., fraction of flow classified as baseflow
BaseflowRecessionK	1/d	Exponential recession constant fitted to master recession curve
First_Recession_Slope	1/d	Steep section of MRC, related to storage that is quickly depleted
Mid_Recession_Slope	1/d	Mid section of MRC, related to water retention capacity of the watershed
EventRR_TotalRR_ratio	-	Ratio between event and total runoff ratio
VariabilityIndex	-	Variability index of flow
Overland flow signatures		
IE_effect	-	Infiltration excess importance (positive coefficients of intensity metrics in a regression equation to predict peak magnitude/ volume)
SE_effect	-	Saturation excess importance (positive coefficients of storage metrics in a regression equation to predict peak magnitude/ volume)
IE_thresh_signif	-	Infiltration excess threshold significance (whether a significant change of slope occurs in a plot of quickflow volume vs. maximum intensity)
SE_thresh_signif	mm/ time- step	Saturation excess threshold significance (whether a significant change of slope occurs in a plot of quickflow volume vs. total precipitation)
IE_thresh	-	Infiltration excess threshold depth (hourly intensity of precipitation needed to produce quickflow in a plot of quickflow volume vs. maximum intensity)
SE_thresh	mm	Saturation excess threshold location (depth of precipitation needed to produce quickflow in a plot of quickflow volume vs. total precipitation)
SE_Slope	mm/ mm	Above-threshold slope in a plot of quickflow volume vs. total precipitation
Storage_thresh_signif	-	Storage threshold significance (whether a significant change of slope occurs in a plot of quickflow volume vs. antecedent precipitation index + total precipitation)
Storage_thresh	mm	Storage threshold location (storage depth needed to produce quickflow in a plot of quickflow volume vs. antecedent precipitation index + total precipitation)

Several of the signatures require watershed-specific parameters, and their values are described in the supporting information, Table S2. We inspected TOSSH errors and warnings to see if any signatures failed, typically due to default parameters being unsuitable for the watershed, or to data errors. Most TOSSH signatures offer a “plot_results” parameter for diagnostic graphical display of signature calculation and values. We used this option for CZOs and for selected watersheds from different regions to visually check the signature calculations (e.g., the fitted recessions). For the CZO watersheds, we made visual checks of the baseflow separation function, that baseflow was adequately separated from quickflow during events. For the event identification function, we checked that the event periods cover major rainfall periods, and those event recession periods include flow peaks occurring immediately after rainfall. We checked that the recession identification function selected all major recession periods and adjusted recession selected parameters if required. For plots where threshold functions were fitted (e.g., to a plot of quickflow against antecedent condition metrics), we checked whether the fit was influenced by a few large or unusual rainstorms.

4 Signature Analysis

We used two approaches to test whether the links between signatures and processes described in McMillan (2020) hold true across multiple watersheds. The first approach used a large sample analysis of signature values across the CAMELS watersheds. For these data, we tested whether signatures related to the same process are correlated, and how signature values are related to climate aridity. The second approach used detailed analyses of signature values in CZO watersheds, to test whether processes inferred from signature values agree with information from watershed-specific literature.

4.1 Large-scale signature and process patterns in CAMELS watersheds

4.1.1 Distribution of signature values

We applied the overland flow and groundwater signature sets across the CAMELS datasets to determine the distributions of values for each signature. This enabled us to classify signature values in terms of their quantile values, i.e., high or low compared to the median for their country, or quasi-globally. This information is valuable for interpretation of the signatures in new watersheds, for example to say whether a recession constant should be considered fast or slow.

4.1.2 Correlation between signature values

Several of the signatures target the same or similar processes, for example multiple signatures indicate high water storage in the watershed. If these signatures represent the same process, we should find correlations between their values. We therefore created a correlation matrix showing rank correlations between each pair of signatures. Spearman rank correlation was used as a nonparametric correlation measure, as relationships between signatures may not be linear. We assessed whether signatures that represent the same or similar processes have high correlations.

4.1.3 Hydro-climate relationship to signature patterns

Climate is a strong control on many signatures (Addor et al., 2018; Knoben et al., 2018). In particular, the aridity index (PET/P) has shown strong (empirical) links to many signatures. We therefore investigated to what extent aridity explains observed signature patterns and whether these patterns are consistent across different countries. We calculated rank correlations and plotted signature values against the aridity index, separated by country. The results will serve as a first assessment of similarity in signature controls across regions, therefore showing how transferable our results might be.

4.2 Signature-process links at Critical Zone Observatories

We based our analysis of signature-process links in CZOs on the summary findings of McMillan (2020; their Table 1). We reorganized their findings into a series of questions about processes in the watershed that could potentially be answered from literature descriptions, and matched the signature values that relate to each question, for example “*Do riparian zones contribute to flow?*” a positive answer implies that there is no rainfall depth threshold before flow occurs, i.e., SE_thresh is close to 0 and/or $SE_thresh_signif > 0.05$. (see Results section 5.2 for the full list of questions, corresponding signature values, and answers). Restructuring the analysis in this way allowed for multiple signatures relating to one process. For each observatory, we collected journal articles describing the watershed processes, and used these to answer the questions. In several cases, the observatories included contrasting sub-watersheds, and process information was collected about each one.

We calculated signature values for all signatures described in Section 3. Where hourly data were available (i.e., all CZOs except Luquillo), we calculated signatures at both hourly and daily timesteps. We compared the impact of timestep choice and noted cases where signature values depend strongly on timestep. We used the distributions of signature values described in Section 4.1.1 to assign a percentile to each value, within the distribution of values across the CAMELS U.S. dataset. We chose to use only the U.S. dataset to quantify percentiles, as we assume that descriptions of processes as being of high or low importance are most likely to implicitly imply a comparison against other U.S. watersheds.

We then placed each question-signature pair into one of three categories: good agreement, mixed agreement, or poor agreement between signature and described process. For CZOs with multiple watersheds with contrasting properties, we assessed whether differences in signatures between these watersheds correspond to known contrasts in hydrologic processes.

5 Results

5.1. Large-scale patterns and distributions

We calculated signature values across each CAMELS dataset. Maps for the four regions and for eight representative signatures are shown in the supporting information (Figures S2-S9).

5.1.1 Distribution of signature values

The distributions of each signature for each CAMELS dataset are shown in Figure 2. These distributions calculated using large samples can be used to assess empirically whether signature

values should be considered high or low. The percentiles of each signature are given in the supporting information, Tables S3 and S4. Some signatures have clearly defined limits (e.g., *BFI*), while others such as watershed storage (*AverageStorage*) have no upper limit. We found that distributions can vary substantially between countries, such that a high signature value in one country might not be considered high in another country (e.g., low values of the *BaseflowRecessionK* in Brazil would be considered average in the U.S., Great Britain, or Australia). Signatures with more uniform and less peaked distributions would be more resilient to uncertainties in the signature value when converting to a percentile of the distribution.

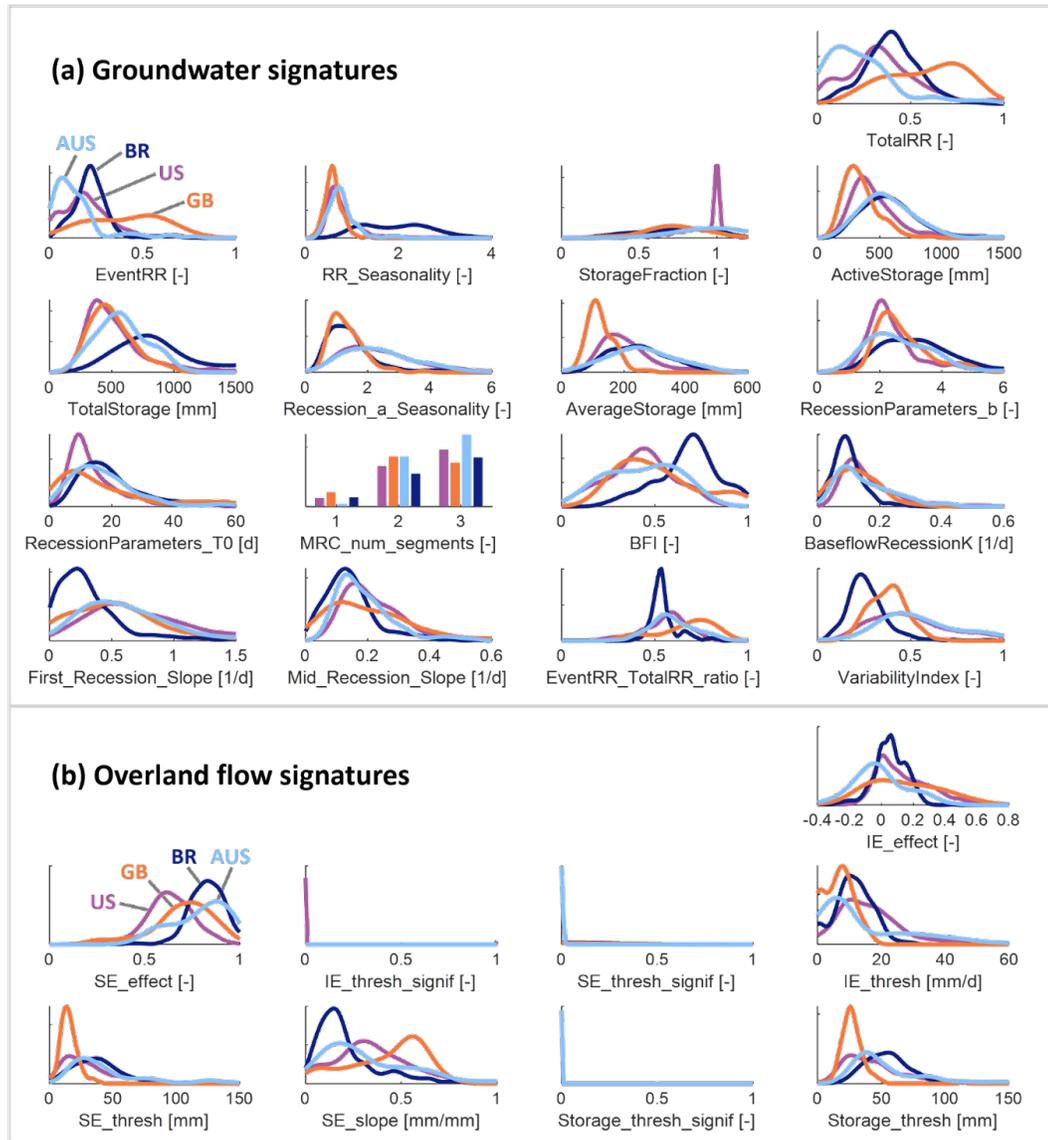


Figure 2. Distributions of (a) groundwater signatures and (b) overland flow signatures, smoothed using a kernel density estimation, except for *MRC_num_segments*, which can only take three values and is thus shown as a bar plot. Plot ranges are adjusted for better visibility.

5.1.2 Correlation between signature values

To test whether signature correlations aligned with physical interpretations of the signatures, we looked for examples where multiple signatures related to the same feature of a flux or store: its magnitude, spatial variation, temporal variation, or response time. The signature correlations are shown in Figure 3.

For baseflow signatures, multiple signatures were available to characterize baseflow magnitude, baseflow response time, and groundwater storage magnitude. High baseflow magnitude and long response time are characterized by high *BFI*, low *BaseflowRecessionK*, low *VariabilityIndex*, low *First_Recession_Slope*, *Mid_Recession_Slope*, and high recession timescale (*RecessionParameters_T0*). These correlations are all correctly represented by the signatures, and provide evidence for the common use of *BFI* as an overarching measure of baseflow importance (Figure 3a). However, the high correlation of baseflow magnitude and response time signatures means that the signatures do not provide a robust method to separate these two aspects of baseflow. Using multiple *BFI* signatures with different time windows would help to resolve this issue (Gnann et al., 2021b).

Several signatures are related to the magnitude of groundwater storage (Table 1 above), including *AverageStorage*, *ActiveStorage*, *TotalStorage*, and *RR_Seasonality*. These signatures are positively correlated, as expected (Figure 3a). However, small values of event runoff ratio (*EventRR*) and its fraction of total runoff ratio (*EventRR_Total_RR_ratio*) are also supposed to signify high storage, and this is not supported by the data. Instead, we found that event runoff ratio and its fraction of *TotalRR* are highly correlated with total runoff ratio, suggesting that they are controlled by losses to evapotranspiration (ET) or deep groundwater as part of the overall water balance. The *StorageFraction* signature is also supposed to relate to storage magnitude but was found to be unreliable, often giving unrealistic values and a poor fit when plotted against the underlying data. This signature was originally developed for a set of 16 watersheds in Luxembourg (Pfister et al., 2017), but modification or generalization of the signature would be needed for it to translate well to other watersheds.

The signature *MRC_num_segments* (number of segments in the master recession curve) has notably low correlations with all other signatures, including *RecessionParameters_b* (nonlinearity in the shape of recession curve) as shown in Figure 3a. However, this is likely due to *MRC_num_segments* being an ordinal signature that can only take values of 1, 2, or 3, and therefore provides less information about the relative values of different watersheds.

For overland flow, we expect negative correlations between saturation excess (*SE_effect*) and infiltration excess (*IE_effect*) and the significance P-value of their thresholds (*IE_thresh_signif*, *SE_thresh_signif*). This holds true for *IE_effect*. For *SE_effect*, the correlation is weaker, showing that precipitation depth can control flow peak and volume but without a threshold in the relationship (Figure 3b). The data also show that threshold size and significance are negatively correlated, showing correctly that the signature algorithm will not identify a large threshold if it is not significant. Although not predicted in advance, we found that all the thresholds (*IE_thresh*, *SE_thresh*, *Storage_thresh*) are strongly positively correlated. High values identify watersheds that require a lot of water to start producing flow, whether this be via infiltration or saturation excess mechanisms.

Overall, we find that expected relationships in baseflow and overland flow signatures hold in general, with a few exceptions as noted above.

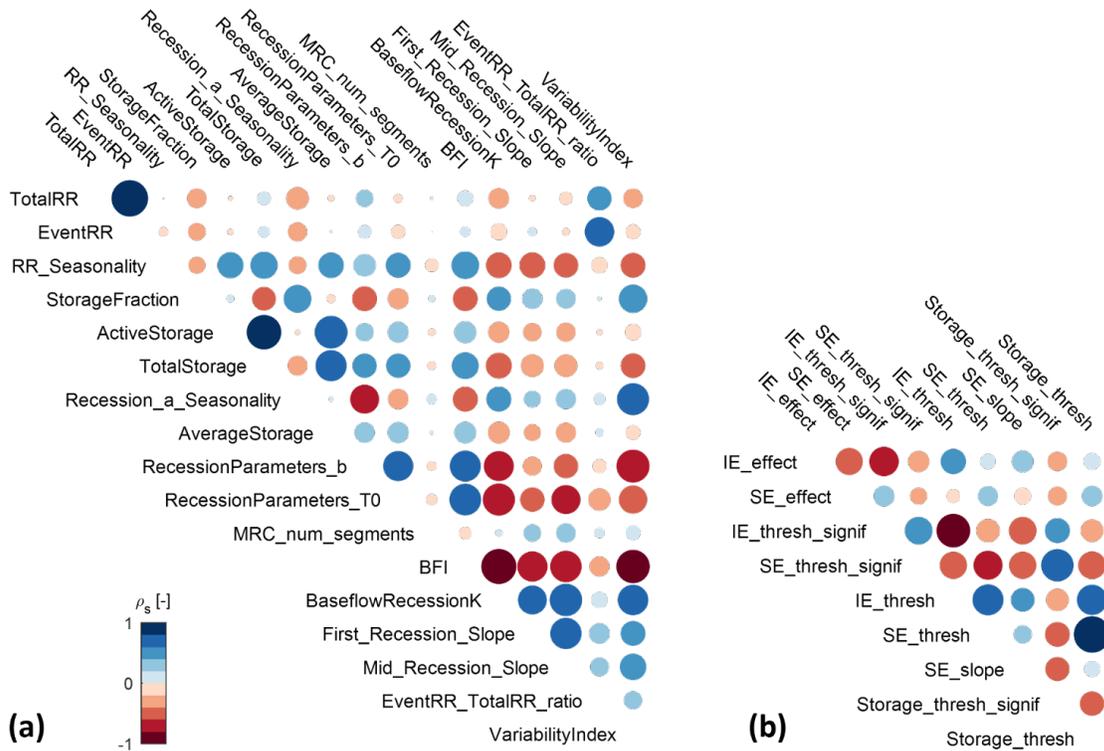


Figure 3. Correlations between (a) groundwater signatures and (b) overland flow signatures for all four CAMELS datasets.

5.1.3 Hydro-climate relationship to signature patterns

Aridity is correlated only with a few signatures (e.g., runoff ratios, see Figure 4a) when all CAMELS watersheds are lumped together (see Figures S10 and S11, and Tables S5 and S6 in the supporting information for all signatures). This changes when countries are investigated separately. For instance, the *IE_effect* signature has a strong correlation with aridity in Great Britain (rank correlation -0.92), while overall it shows only a very weak correlation (rank correlation -0.24), see Figure 4b. Sometimes, the relationships even have opposite signs, as is the case for the BFI in Great Britain and Australia, see Figure 4c. As we expected, these results show that signatures are not solely controlled by climate, but also by other watershed characteristics (e.g., soils, geology). An understanding of how climate and landscape characteristics interact will be essential in predicting signature values. Our results demonstrate that relationships between climate characteristics and signatures from a single region should not be assumed to hold in other regions, and show the benefits of using multi-continent datasets to understand drivers of signature patterns.

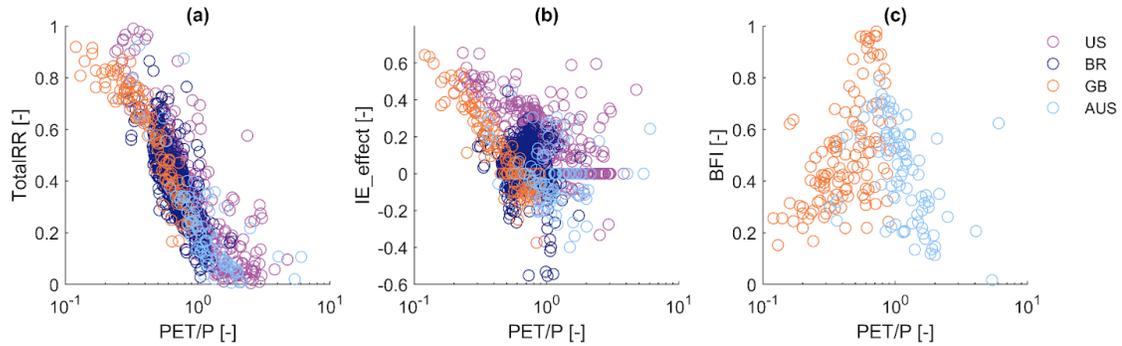


Figure 4. Relationships between **(a)** TotalRR and aridity (PET/P) for all CAMELS countries, **(b)** between IE_effect and aridity for all CAMELS countries, and **(c)** between BFI and aridity for Great Britain and Australia.

5.2 Signature-process links at Critical Zone Observatories

As described in the Methods section, we collected literature from each CZO to answer each process question, and calculated the values and percentiles of each signature matched to that question. This information allowed us to describe the signature-process agreement in text, and ascribe good, partial, or poor agreement between each pair. For these well-studied watersheds, most of the process questions could be answered by searching the literature, but there were some gaps (white cells in the table). A full spreadsheet showing all signature values, percentiles, descriptions of each process, and key references for each CZO watershed is given in the Github repository (see data availability section). A summary showing signature-process agreement is shown here (Figure 5).

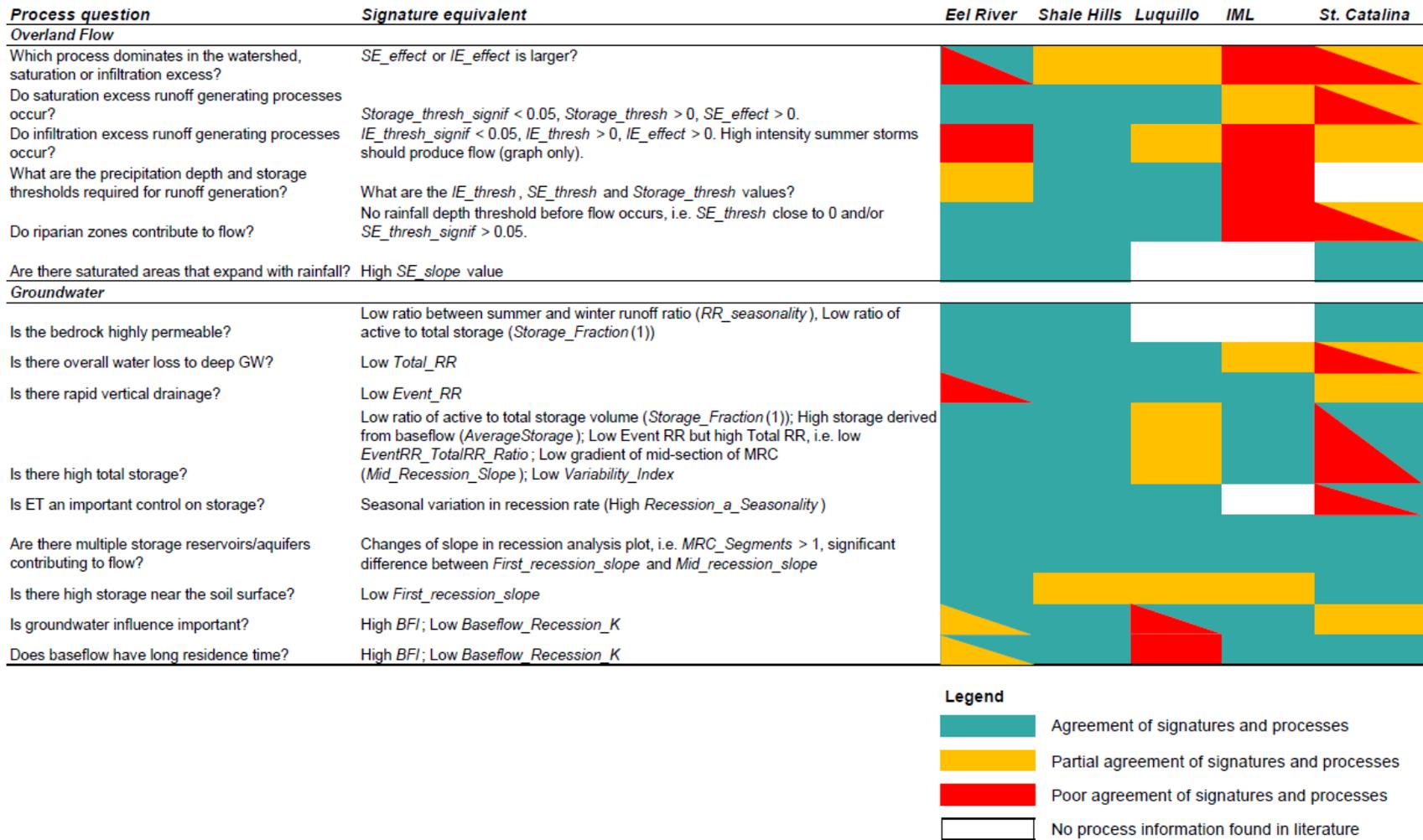


Figure 5. Questions derived from McMillan (2020), used to compare signature values and process knowledge. For each CZO watershed, cells are color-coded for process knowledge and signature values agreement (blue-green), partial agreement (yellow), or poor agreement (red). Where results differ substantially between signatures and/or subwatersheds, split colors are used. White cells imply that we found insufficient literature to answer the question.

In general, there was good agreement between the processes interpreted from literature and the corresponding signatures (blue-green in Figure 5). Overall, groundwater signatures (71% agreement, 20% partial agreement, 8% poor agreement) were more reliable than overland flow signatures (46% agreement, 28% partial agreement, 26% poor agreement).

Good matches between signatures and processes occur in Eel River, Shale Hills, and Luquillo CZOs. Eel River in Northern California has two contrasting sub-watersheds, Elder Creek with high groundwater storage, and Dry Creek with low groundwater storage and frequent saturation excess flow. Groundwater signatures accurately represent the contrasts in storage and seasonality. However, SE flow in Dry Creek is incorrectly identified as IE at daily timescale, and BFI is moderate (not high as expected) in Elder Creek, representing a compromise between the Mediterranean climate that favors low baseflow and high storage that favors high baseflow. Shale Hills in Pennsylvania lies on sedimentary geology in the Appalachian Mountains. In wet conditions, rising water tables generate interflow with a transmissivity-feedback mechanism. Deep groundwater and surface flow are smaller components, although saturated riparian areas generate flow. Signatures and processes agree across almost all overland flow and groundwater processes, predicting low storage and BFI, fast recessions, a storage threshold for flow, seasonal ET influence, and low/moderate surface flow. Luquillo watersheds in Puerto Rico comprise tropical, montane forest, with Rio Icaros (granitoid) and Rio Mameyes (volcaniclastic) on contrasting geologies. Fast/shallow processes dominate despite deep soils, with event water flowing as a perched water table. Signatures and processes mostly agree, predicting fast processes (high event runoff coefficient and steep initial recessions), a storage threshold for flow, low seasonality, and no clear IE or SE dominance. Granitoid Icaros correctly has higher BFI and lower Recession K than Mameyes, but for both rivers the recession parameters incorrectly merge fast event processes with a small but sustained baseflow component.

Poorer matches between signatures and processes occur in IML and Santa Catalina CZOs. The IML Upper Sangamon watershed in Illinois is in row-crop agriculture with tile drains. SE and IE flow are both reported in the literature, but with IE dominant. Groundwater rises quickly after events and runs off via tile drains. Groundwater signatures and processes mostly agree, with low event runoff ratio suggesting drainage to groundwater, and low storage signatures suggesting that this groundwater drains quickly to the stream. However, overland flow processes are incorrectly identified as dominated by SE, with IE found not significant although it is known to occur. The Santa Catalina watersheds in Northern Arizona are arid, mid to high-elevation sites. Heavy summer monsoon storms produce overland and near-surface flows dominated by event and soil water. IE dominates, with some near-stream SE. Streams gain some water from deep/regional groundwater. Signatures and processes often disagree; signatures incorrectly suggest that saturation and storage processes dominate, and very low runoff ratios make interpretation difficult. However, moderate BFI and recession K agree with the limited but important groundwater contribution.

6 Discussion

In this paper, we applied hydrological signatures and assessed their process interpretations in a diverse set of basins, in many cases well outside of the hydroclimatic regimes for which the signatures were designed. Therefore, we gained many useful insights into signature use.

6.1 Benefits of calculating signature distributions

To interpret signature values as ‘high’ or ‘low’, it was essential to know the distributions of signature values (Figure 2). For example, *SE_effect* values (saturation excess importance) are consistently higher than *IE_effect* values (infiltration excess importance), so these values are interpreted differently. It was therefore useful to present the signature value as a percentile of the national distribution rather than an absolute value. We recommend a comparison with regional rather than global distributions, as we hypothesize that authors implicitly compare processes within the same region. For example, a low *BFI* in Brazil might be considered a high *BFI* in the U.S.

Analyzing signature values on a national scale was also useful when understanding correlations of signatures with aridity. We found examples where a correlation that occurs for the combined set of four CAMELS data sets does not hold for individual countries. Strong but diverging correlations (e.g., *BFI*, see Figure 4c) might point at relationships that are not causal. For instance, the most productive aquifers in Great Britain happen to be in the least humid places, so a correlation between aridity and *BFI* here might be a coincidence, and this could be the reason why it does not hold for other countries. Understanding such regional differences in relationships will be essential for prediction of signatures in ungauged basins.

6.2 Signature robustness

Large sample signature calculations pose several challenges. Some signatures are straightforward to calculate (e.g., *TotalRR*), or have been widely used (e.g., *BFI*) and are relatively robust. Some watershed types may prove more susceptible to uncertainties or difficulties in signature calculation, such as leaky watersheds that cause errors in estimated ET (Wlostowski et al., 2020). Signatures that have only been used in a few small scale studies, or are sensitive to parameter selections (e.g., recessions, see Dralle et al., 2017; Stoelzle et al., 2013) are less robust for large samples. Their results might not be reliable, even though the values might be within a realistic range. For example, the monsoonal climate in some parts of Brazil leads to a distinct seasonal flow regime, with many short recessions during the wet season and a long recession during the dry season (Figure 6). This is not picked up by the recession signatures we used, as they return a single (average) storage-discharge relationship. It is thus important to check visually whether the signature results are reasonable, and to test signatures when transferring them to other scales or other places. It is then possible to tailor the signatures to certain regions, e.g., by dividing the time series according to season (Euser et al., 2013).

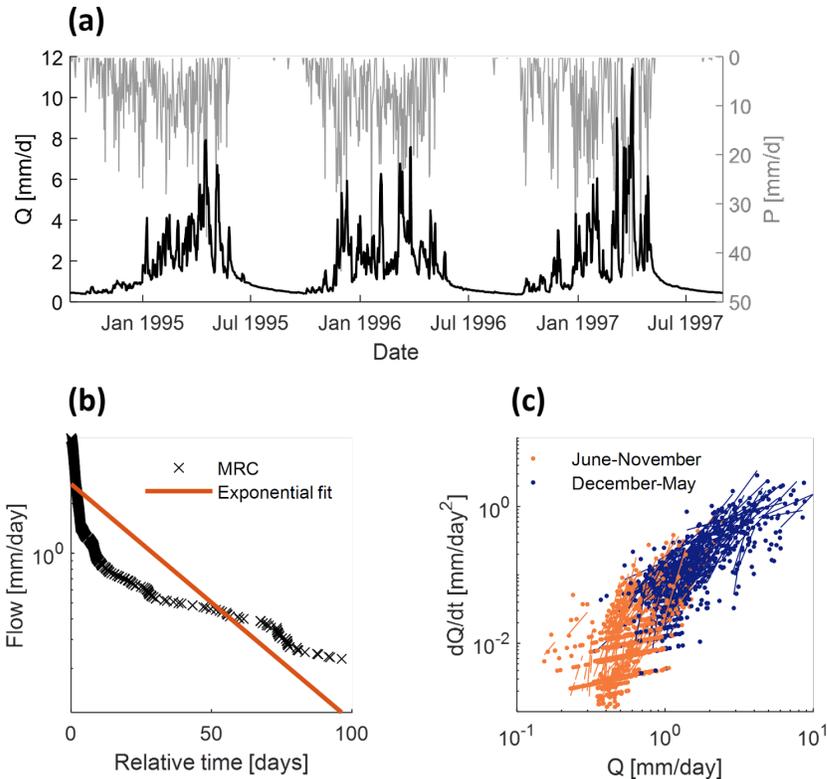


Figure 6. (a) Hydrograph of a watershed in Brazil that shows very seasonal (monsoonal) precipitation and thus streamflow. (b) The corresponding master recession curve flattens out for late recessions, indicating the transition from wet season recessions to dry season recessions. (c) The same can be seen from the dQ/dt plot where dry season recessions are systematically steeper.

Although we expected hourly data to provide more accurate estimates of event characteristics and recession dynamics in small watersheds, we found that working with hourly data required a hands-on approach to prevent errors. This included changing recession selection tolerance due to diurnal flow fluctuations, and filling gaps in timeseries; for example, USGS flow data is usually infilled at the daily timescale, but not for hourly data. Such interventions cause hourly signature values to be more uncertain, trading off accuracy and data processing time. We suggest comparing hourly and daily signature values, and identifying reasons for significant differences. In the arid Santa Catalina CZO, hourly data gave poor results that did not match literature information on processes. We do however recommend hourly data for identifying overland flow processes, as it produced results that better matched field observations, and were often substantially different from daily results.

6.3 Lessons from comparisons of signatures and processes for CZOs

We found some challenges in the CZO watersheds when comparing signature values to process descriptions. It could be difficult to obtain standardized process data, such as depth to bedrock which was sometimes quantified differently by different authors, even in the same watershed. We sometimes found conflicting information, such as in the Intensively Managed Landscapes CZO, where saturation excess flow was said to occur, but the water table was said to be low due to tile drains. Such conflicts could be due to differences in exact location, or in wetness conditions at the

time of observation, and illustrate the difficulties in summarizing complex understanding of the landscape. Similarly, some signature values failed to capture the full insights of field studies, such as in Shale Hills CZO that has a known discharge of old (20-30yr) water to the stream, but which comprises only a small percentage of flow. Therefore, the BFI signature value is low and downplays this discharge.

Some processes were less reliable in their match to signature values across multiple CZOs. In particular, IE and SE processes (first three rows in Figure 5) were not well differentiated. For example, SE is known to dominate in Dry Creek in the Eel River CZO, but signatures show IE; whereas IE is known to dominate in IML CZO, but signatures show SE. Additionally, watersheds differed in their reliability, for example, Shale Hills showed high reliability across all signatures with no disagreements. In the arid but high-elevation Santa Catalina CZO, the reliability of signatures based on events and recession periods was reduced, because only a small number of storms produced flow, and some of these were impacted by snow. The IML CZO, where the hydrology reflects significant human impacts in cropped areas, showed low reliability for overland flow signatures. The low reliability could be due to high variability in processes between impacted and non-impacted areas of the watershed, as overland flow signatures face a scale conflict between location-specific observations of flow, and signature values that reflect integrated watershed response. Low reliability could also be due to human impacts such as tile drains. Signatures are known to be modified by human activities, such as baseflow index being affected by groundwater abstraction and effluent discharges to rivers (Bloomfield et al., 2021), but were not originally designed for use in human-impacted watersheds.

6.4 Comparison with previous studies

It is useful to compare our results with previous studies that related process descriptions to signature values. For overland flow processes, Wu et al. (2021) identified infiltration and saturation excess using Spearman correlations between event runoff ratios and rainfall intensity, rainfall volume and rainfall storage. They found few watersheds with dominant infiltration excess, agreeing with previous findings that IE flow is rare in the U.S. (Buchanan et al., 2018; Wolock, 2003). However, there are substantial differences in spatial patterns of IE between our study and these previous studies, and among the previous studies. One explanation for high uncertainty is that overland flow signatures are sensitive to calculation methods, particularly whether hourly or daily rainfall intensity is used, and require multiple choices including baseflow separation, event definition, and storage calculation methods. Our *IE_effect* and *SE_effect* signatures are based on a study by Estrany et al., (2010) in a Mediterranean watershed, but may not function correctly in other climates, as also evidenced by the unexpected positive correlation between wetness and *IE_effect* in daily CAMELS-GB data. Our *IE_thresh* and *SE_thresh* signatures were more consistent with previous studies and process knowledge, particularly when using hourly rainfall intensity. They showed positive infiltration excess thresholds in arid and Southeastern U.S., where infiltration excess is expected to occur, and positive saturation excess thresholds in most of the U.S. except the arid West and in the North East where antecedent conditions may outweigh event volume. In summary, *IE_effect* and *SE_effect* signatures are not reliable, and future work is needed to design and test signatures that better differentiate these processes.

For groundwater processes, we can compare our results with those of Wlostowski et al. (2020), who studied how critical zone architecture controls signature values. Our results agree with theirs in finding that baseflow and storage signatures are controlled not by depth to bedrock but rather

by properties and structures of the soil. For example, expert observations of whether shallow interflow and return flow processes occur were more likely to match signature values than a simple depth to bedrock value. An example occurs in the Luquillo CZO, where depth to bedrock at the Rio Mameyes site is 30-40m, but streamflow dynamics are dominated by rapid delivery of event water to the stream by fast, shallow runoff processes including lateral macropore flow. Wlostowski et al. (2020) further agreed with our findings in noting a clear influence of tile drains in signatures for the IML CZO.

6.5 Recommendations for signature choice

In this section we record signature-specific conclusions from our study, in particular whether signatures could be robustly calculated across large samples of watersheds, whether signatures related correctly to process interpretations, and whether signatures relied on any watershed-specific fitting parameters. Fitting parameters that affect multiple signatures, and would ideally be checked visually against the flow timeseries, are those that control event selection and recession selection. The recommendations are summarized in Table 2.

Table 2. Summary of findings for groundwater and overland flow signatures.

Signature	Recommendation/Comments
Groundwater and baseflow signatures	
TotalRR	Strongly related to aridity across all CAMELS datasets; describes climate more than hydrology. Easy and robust to calculate.
EventRR	Highly correlated to Total RR, interpretation therefore relates to total water balance rather than to watershed storage as previously thought. Relies on multiple parameters to control event selection that may be region-specific.
RR_Seasonality	Agreed well with process interpretations in CZO watersheds, related to bedrock permeability and watershed storage size. Correlates well to other storage magnitude signatures. Relies on recession selection parameters.
StorageFraction (incl. ActiveStorage and TotalStorage)	Unreliable signature for large samples as seen in unrealistic values and poor fit in plotting (shown in the supporting information Figure S1).
Recession_a_Seasonality	Agreed well with process interpretations in CZO watersheds, relates to ET influence on storage. Relies on recession selection parameters.
AverageStorage	More reliable than the StorageFraction above, recommended when estimates of storage are needed. Good agreement with process knowledge in CZO watersheds.
RecessionParameters (b, T0)	Established signatures with theoretical link to watershed storage-discharge relationship, highly correlated to other signatures of baseflow magnitude and response time. Does not distinguish short event recessions from longer dry season recessions. Relies on recession selection parameters.
MRC_num_segments	Useful signature to identify complexity of recession shapes, robust across a wide range of recession characteristics. Relies on recession selection parameters. Where recessions have multiple segments with different slopes, the BaseflowRecessionK signature may be unreliable.
BFI	Reliable signature that correlates strongly with most other baseflow signatures. However, note that BFI integrates multiple aspects of baseflow

	(volume of baseflow, response time, multiple baseflow sources or pathways), which could not be distinguished based on a single BFI value.
BaseflowRecessionK	Usually reliable, correlates well with other baseflow signatures. However, when using high BFI and Low Baseflow_Recession_K as indicators of long GW residence times, it does not properly differentiate between a little baseflow with long residence time or a moderate amount of baseflow with moderate residence times. We advise a visual check of fit when MRC_num_segments is > 1.
First_Recession_Slope	MRC fitting was robust across varied recession shapes. A low first slope is supposed to indicate high storage near the soil surface, i.e., the fastest flow path is delayed due to this storage. However, it is not easy to determine whether this matches with soil profile observations that typically record soil texture and/or importance of shallow flow processes. Relies on recession selection parameters.
Mid_Recession_Slope	MRC fitting was robust across varied recession shapes, correlates well with other baseflow signatures. Relies on recession selection parameters.
EventRR_TotalRR_ratio	Event and Total RR are highly correlated, making this ratio more uncertain. Showed reasonable match to process interpretations in most CZOs, except for arid Santa Catalina watershed. Relies on multiple parameters to control event selection.
VariabilityIndex	Showed a moderate fit to storage information in CZO watersheds and easier to calculate than AverageStorage signature although the latter is preferred for storage estimates.
Overland flow signatures	
IE_effect	IE_effect is not a reliable signature. Watersheds where IE or SE dominates can be incorrectly identified by IE_effect and SE_effect values. Hourly data produces a closer match to process knowledge. Relies on event selection parameters.
SE_effect	SE_effect is not a reliable signature. Not strongly related to the threshold signatures; may show control of rainfall depth on flow, independent of the existence of a threshold. SE_effect is above 0.5 for most watersheds, so need the percentile to quantify high/low values. Hourly data produces a closer match to process knowledge. Relies on event selection parameters.
IE_thresh_signif	Strongly negatively correlated to IE_thresh: i.e., the code will generally not identify a threshold if it is not significant, but useful to confirm threshold existence. Relies on event selection parameters.
SE_thresh_signif	Strongly negatively correlated to SE_thresh: i.e., the code will generally not identify a threshold if it is not significant, but useful to confirm threshold existence. Relies on event selection parameters.
IE_thresh, SE_thresh, Storage_thresh	These signatures are more reliable than IE/SE_effect: large, significant thresholds suggest that these processes are important. All the thresholds (IE, SE, storage thresh) are strongly positively correlated. i.e., this identifies watersheds that require a lot of water to start producing flow. The very strong correlation between Storage and SE thresholds shows a difficulty in separating the impacts of pre-event storage and event depth. To achieve this, a longer storage memory than Mosleys (Mosley, 1979) 30-day API used in

	the Storage_thresh signature may be required. Hourly rainfall data is preferred for IE_thresh. Relies on event selection parameters.
SE_slope	Agreed well with process interpretations in CZO watersheds, relates to expansion of saturated areas with event precipitation. Relies on event selection parameters.
Storage_thresh_signif	Moderately negatively correlated to SE_thresh: i.e., the code will usually not identify a threshold if it is not significant. Useful to confirm threshold existence, but see comments on Storage_thresh above. Relies on event selection parameters.

7 Conclusions

This study tested whether relationships between signatures and processes developed from experimental watershed studies hold true when applied over large scales and diverse hydroclimates. The relationships were tested using two types of data: large sample CAMELS datasets from four regions, and detailed information from five CZO watersheds in the U.S. We note that when single signature values are used to summarize complex watershed responses, they might represent a compromise value between climate and process effects, or between multiple processes (e.g., fast and slow recession processes, or spatially variable overland flow). This compromise demonstrates the difficulty of summarizing processes using quantitative values, without losing some information.

We identified a small number of signatures that were not reliable (*SE_effect*, *IE_effect*, *StorageFraction*) or had different interpretations than expected (*EventRR*). We made recommendations for adapting some signatures to better differentiate between processes (using multiple BFI timescales to separate baseflow magnitude and response time, calculating *Storage_thresh* with longer memory to differentiate from *SE_thresh*). Overall, the results showed that most signature patterns agreed with process interpretations, with groundwater and baseflow signatures being more reliable than overland flow signatures. Based on the CZO watershed results, signature-process relationships were most reliable in humid and Mediterranean-climate watersheds, and less reliable in arid and human-impacted watersheds. This difference reflects the history of signature development which has been concentrated in natural, humid basins, and points to scope for future signature development in a wider range of watersheds.

Open Research

Data and Code Availability

The CAMELS U.S. dataset is available at <https://dx.doi.org/10.5065/D6MW2F4D> (Addor et al., 2017; Newman et al., 2015). The hourly rainfall dataset corresponding to the CAMELS U.S. locations is available at <https://doi.org/10.5281/zenodo.4072700> (Gauch et al., 2020, 2021). The CAMELS Great Britain dataset is available at <https://doi.org/10.5285/8344e4f3-d2ea-44f5-8afa-86d2987543a9> (Coxon et al., 2020). The CAMELS Australia dataset is available at <https://doi.pangaea.de/10.1594/PANGAEA.921850> (Fowler et al., 2021). The CAMELS Brazil dataset is available at <https://zenodo.org/record/3964745> (Chagas et al., 2020). CZO data products are available at <https://doi.org/10.4211/hs.29e2ec85770b42c881ef0750696463e5> (Wlostowski et al., 2021). The TOSSH toolbox (Gnann et al., 2021a) used to calculate hydrologic signatures is

available at <https://github.com/TOSSHtoolbox/TOSSH>. The code used to load the CAMELS data into Matlab is available at https://github.com/SebastianGnann/CAMELS_Matlab. The code to reproduce our analysis, and a full spreadsheet showing all signature values, percentiles, descriptions of each process and key references for each CZO watershed, are available at <https://github.com/SebastianGnann/LargeScaleSigs>.

References

- Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N., & Clark, M. P. (2018). A ranking of hydrological signatures based on their predictability in space. *Water Resources Research*, *54*(11), 8792–8812. <https://doi.org/10.1029/2018WR022606>
- Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences (HESS)*, *21*(10), 5293–5313. <https://doi.org/10.5194/hess-21-5293-2017>
- Beighley, R. E., Dunne, T., & Melack, J. M. (2005). Understanding and modeling basin hydrology: Interpreting the hydrogeological signature. *Hydrological Processes*, *19*(7), 1333–1353. <https://doi.org/10.1002/hyp.5567>
- Black, P. E. (1997). Watershed functions. *Journal of the American Water Resources Association*, *33*(1), 1–11. <https://doi.org/10.1111/j.1752-1688.1997.tb04077.x>
- Bloomfield, J. P., Gong, M., Marchant, B. P., Coxon, G., & Addor, N. (2021). How is Baseflow Index (BFI) impacted by water resource management practices? *Hydrology and Earth System Sciences*, *25*(10), 5355–5379. <https://doi.org/10.5194/hess-25-5355-2021>
- Buchanan, B., Auerbach, D. A., Knighton, J., Evensen, D., Fuka, D. R., Easton, Z., Wiczorek, M., Archibald, J. A., McWilliams, B., & Walter, T. (2018). Estimating dominant runoff modes across the conterminous United States. *Hydrological Processes*, *32*(26), 3881–3890. <https://doi.org/10.1002/hyp.13296>
- Chagas, V. B., Chaffe, P. L., Addor, N., Fan, F. M., Fleischmann, A. S., Paiva, R. C., & Siqueira, V. A. (2020). CAMELS-BR: Hydrometeorological time series and landscape attributes for 897 catchments in Brazil. *Earth System Science Data*, *12*(3), 2075–2096. <https://doi.org/10.5194/essd-12-2075-2020>
- Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., & Gochis, D. J. (2015). A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resources Research*, *51*(4), 2515–2542. <https://doi.org/10.1002/2015WR017200>
- Coxon, G., Addor, N., Bloomfield, J. P., Freer, J., Fry, M., Hannaford, J., Howden, N. J., Lane, R., Lewis, M., & Robinson, E. L. (2020). CAMELS-GB: Hydrometeorological time series and landscape attributes for 671 catchments in Great Britain. *Earth System Science Data*, *12*(4), 2459–2483. <https://doi.org/10.5194/essd-12-2459-2020>
- Dralle, D. N., Karst, N. J., Charalampous, K., Veenstra, A., & Thompson, S. E. (2017). Event-scale power law recession analysis: Quantifying methodological uncertainty. *Hydrology and Earth System Sciences*, *21*(1), 65. <https://doi.org/10.5194/hess-2016-341>
- Dunne, T. (1978). Field studies of hillslope flow processes. In *M. J. Kirkby (Ed.), Hillslope hydrology* (pp. 227–293). John Wiley & Sons.
- Estrany, J., Garcia, C., & Batalla, R. J. (2010). Hydrological response of a small mediterranean agricultural catchment. *Journal of Hydrology*, *380*(1–2), 180–190. <https://doi.org/10.1016/j.jhydrol.2009.10.035>

- Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., & Savenije, H. H. G. (2013). A framework to assess the realism of model structures using hydrological signatures. *Hydrology and Earth System Sciences*, *17*(5), 1893–1912. <https://doi.org/10.5194/hess-17-1893-2013>
- Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., & Grant, G. (2019). Hillslope hydrology in global change research and Earth system modeling. *Water Resources Research*, *55*(2), 1737–1772. <https://doi.org/10.1029/2018wr023903>
- Fowler, K. J., Acharya, S. C., Addor, N., Chou, C., & Peel, M. C. (2021). CAMELS-AUS: Hydrometeorological time series and landscape attributes for 222 catchments in Australia. *Earth System Science Data Discussions*, 1–30. <https://doi.org/10.5194/essd-13-3847-2021>
- Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., & Hochreiter, S. (2021). Rainfall–runoff prediction at multiple timescales with a single Long Short-Term Memory network. *Hydrology and Earth System Sciences*, *25*(4), 2045–2062. <https://doi.org/10.5194/hess-25-2045-2021>
- Gauch, Martin, Kratzert, Frederik, Klotz, Daniel, Nearing, Grey, Lin, Jimmy, & Hochreiter, Sepp. (2020). Models and Predictions for "Rainfall-Runoff Prediction at Multiple Timescales with a Single Long Short-Term Memory Network" [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.4095485>
- Gnann, S. J., Coxon, G., Woods, R. A., Howden, N. J., & McMillan, H. K. (2021a). TOSSH: A Toolbox for Streamflow Signatures in Hydrology. *Environmental Modelling & Software*, *138*, 104983. <https://doi.org/10.1016/j.envsoft.2021.104983>
- Gnann, S. J., McMillan, H. K., Woods, R. A., & Howden, N. J. K. (2021b). Including Regional Knowledge Improves Baseflow Signature Predictions in Large Sample Hydrology. *Water Resources Research*, *57*(2), e2020WR028354. <https://doi.org/10.1029/2020WR028354>
- Grayson, R., & Blöschl, G. (2001). Summary of pattern comparison and concluding remarks. *Spatial Patterns in Catchment Hydrology: Observations and Modelling.*, 355–396.
- Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., & Andréassian, V. (2014). Large-sample hydrology: A need to balance depth with breadth. *Hydrology and Earth System Sciences*, *18*(2), 463–477. <https://doi.org/10.5194/hess-18-463-2014>
- Hellie, F., Peschke, G., Seidler, C., & Niedel, D. (2002). Process-oriented subdivision of basins to improve the preprocessing of distributed precipitation-runoff-models. *Interdisciplinary Approaches in Small Catchment Hydrology: Monitoring and Research*, 137.
- Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2018). A Quantitative Hydrological Climate Classification Evaluated With Independent Streamflow Data. *Water Resources Research*, *54*(7), 5088–5109. <https://doi.org/10.1029/2018WR022913>
- Lundquist, J. D., Cayan, D. R., & Dettinger, M. D. (2004). Spring onset in the Sierra Nevada: When is snowmelt independent of elevation? *Journal of Hydrometeorology*, *5*(2), 327–342. [https://doi.org/10.1175/1525-7541\(2004\)005%3C0327:SOITSN%3E2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005%3C0327:SOITSN%3E2.0.CO;2)
- Markstrom, S. L., Hay, L. E., & Clark, M. P. (2016). Towards simplification of hydrologic modeling: Identification of dominant processes. *Hydrology and Earth System Sciences*, *20*(11), 4655–4671. <https://doi.org/10.5194/hess-20-4655-2016>
- McDonnell, J. J., & Woods, R. (2004). On the need for catchment classification. *Journal of Hydrology*, *299*(1), 2–3. <https://doi.org/10.1016/j.jhydrol.2004.09.003>
- McMillan, H. (2020). Linking hydrologic signatures to hydrologic processes: A review. *Hydrological Processes*, *34*(6), 1393–1409. <https://doi.org/10.1002/hyp.13632>
- McMillan, H. K. (2021). A review of hydrologic signatures and their applications. *Wiley Interdisciplinary Reviews: Water*, *8*(1), e1499. <https://doi.org/10.1002/wat2.1499>
- Mosley, M. P. (1979). Streamflow generation in a forested watershed, New Zealand. *Water Resources Research*, *15*(4), 795–806. <https://doi.org/10.1029/wr015i004p00795>
- Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., & Arnold, J. R. (2015). Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of

- regional variability in hydrologic model performance. *Hydrology and Earth System Sciences*, 19(1), 209–223. <https://doi.org/10.5194/hess-19-209-2015>
- Peschke, G., Etzenberg, C., Töpfer, J., Zimmermann, S., & Müller, G. (1999). Runoff generation regionalization: Analysis and a possible approach to a solution. *IAHS-AISH Publication*, 147–156.
- Pfister, L., Martínez-Carreras, N., Hissler, C., Klaus, J., Carrer, G. E., Stewart, M. K., & McDonnell, J. J. (2017). Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments. *Hydrological Processes*, 31(10), 1828–1845. <https://doi.org/10.1002/hyp.11134>
- Scherrer, S., & Naef, F. (2003). A decision scheme to indicate dominant hydrological flow processes on temperate grassland. *Hydrological Processes*, 17(2), 391–401. <https://doi.org/10.1002/hyp.1131>
- Sivakumar, B. (2008). Dominant processes concept, model simplification and classification framework in catchment hydrology. *Stochastic Environmental Research and Risk Assessment*, 22(6), 737–748. <https://doi.org/10.1007/s00477-007-0183-5>
- Sivapalan, M. (2006). Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the Catchment Scale. In *Encyclopedia of Hydrological Sciences*. John Wiley & Sons. <https://doi.org/10.1002/0470848944.hsa012>
- Stoelzle, M., Stahl, K., & Weiler, M. (2013). Are streamflow recession characteristics really characteristic? *Hydrology and Earth System Sciences*, 17(2), 817–828. <https://doi.org/10.5194/hess-17-817-2013>
- Wagener, T., Sivapalan, M., Troch, P., & Woods, R. (2007). Catchment classification and hydrologic similarity. *Geography Compass*, 1(4), 901–931. <https://doi.org/10.1111/j.1749-8198.2007.00039.x>
- Wlostowski, A. N., Molotch, N., Anderson, S. P., Brantley, S. L., Chorover, J., Dralle, D., Kumar, P., Li, L., Lohse, K. A., & Mallard, J. M. (2020). Signatures of Hydrologic Function Across the Critical Zone Observatory Network. *Water Resources Research*, 57(3). <https://doi.org/10.1029/2019wr026635>
- Wolock, D. M. (2003). *Infiltration-excess overland flow estimated by TOPMODEL for the conterminous United States*. US Geological Survey. Retrieved from <https://pubs.er.usgs.gov/publication/ofr03310>
- Wu, S., Zhao, J., Wang, H., & Sivapalan, M. (2021). Regional patterns and physical controls of streamflow generation across the conterminous United States. *Water Resources Research*, 57(6). <https://doi.org/10.1029/2020wr028086>
- Zotarelli, L., Dukes, M. D., Romero, C. C., Migliaccio, K. W., & Morgan, K. T. (2010). Step by step calculation of the Penman-Monteith Evapotranspiration (FAO-56 Method). *Institute of Food and Agricultural Sciences. University of Florida*.

Large Scale Evaluation of Relationships between Hydrological Signatures and Processes

H. McMillan¹, S. J. Gnann², and R. Araki¹

¹Department of Geography, San Diego State University, San Diego, CA, USA

²Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany

Contents of this file

Figures S1 to S11

Tables S1 to S6

Introduction

This supporting information contains Figures and Tables that were not included in the main manuscript. Figures and Tables present either additional information or full results of the analysis.

Figure S1 illustrates the uncertainties in calculating StorageFraction signatures. Figures S2 to S9 present maps of eight selected signatures (AverageStorage, EventRR, RecessionParameters_b, BaseflowRecessionK, IE_effect, SE_effect, IE_thresh, and SE_thresh). Figures S10 and S11 present relationships between signatures and aridity.

Tables S1 and S2 provide additional information on signature selection and calculation methods. Tables S3 and S4 show signature distributions converted to percentiles. Tables S5 and S6 show rank correlations of signature values and aridity by country.

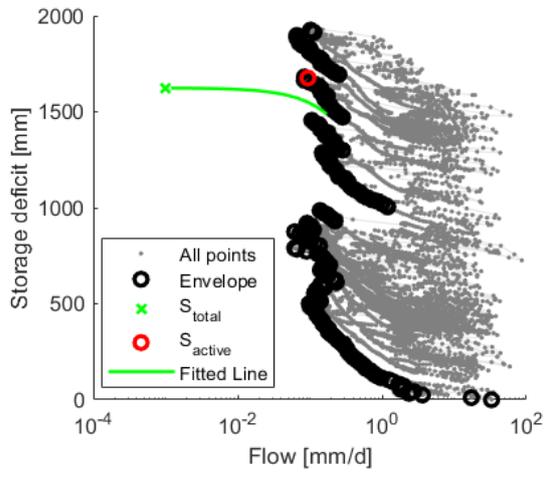


Figure S1. Plot showing StorageFraction signature calculation.

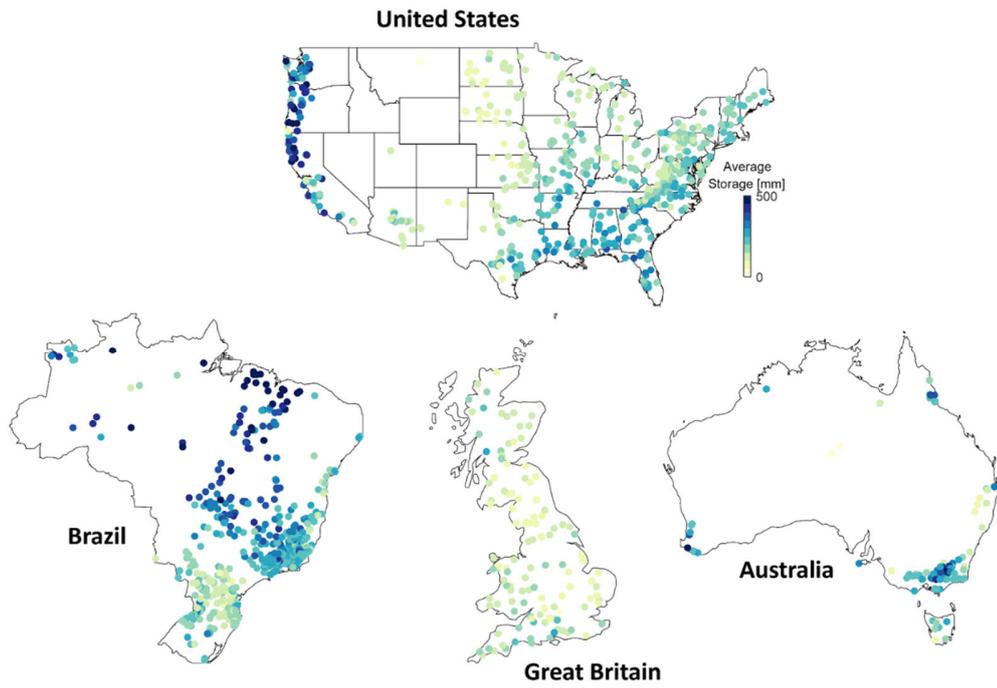


Figure S2. Maps of AverageStorage. Note that the maps of the countries are not to the same scale.

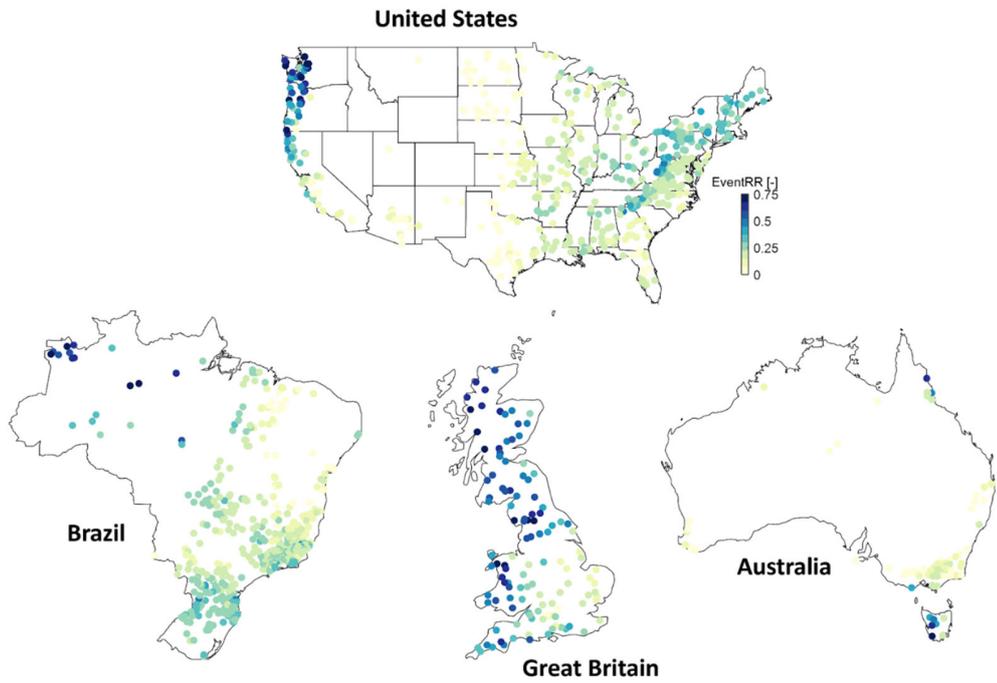


Figure S3. Maps of EventRR. Note that the maps of the countries are not to the same scale.

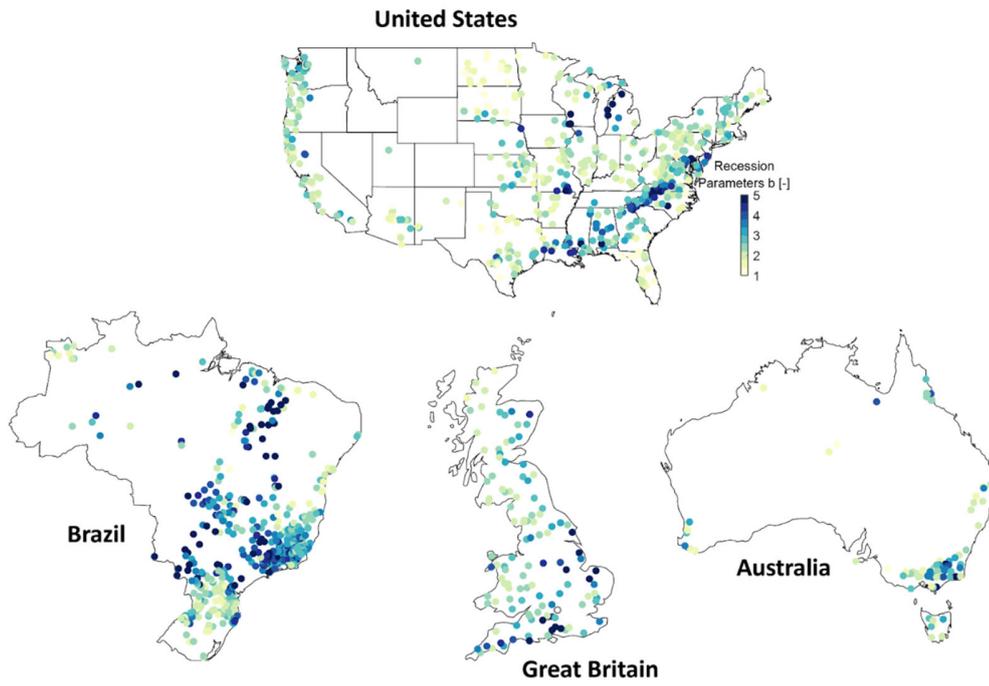


Figure S4. Maps of RecessionParameters_b. Note that the maps of the countries are not to the same scale.

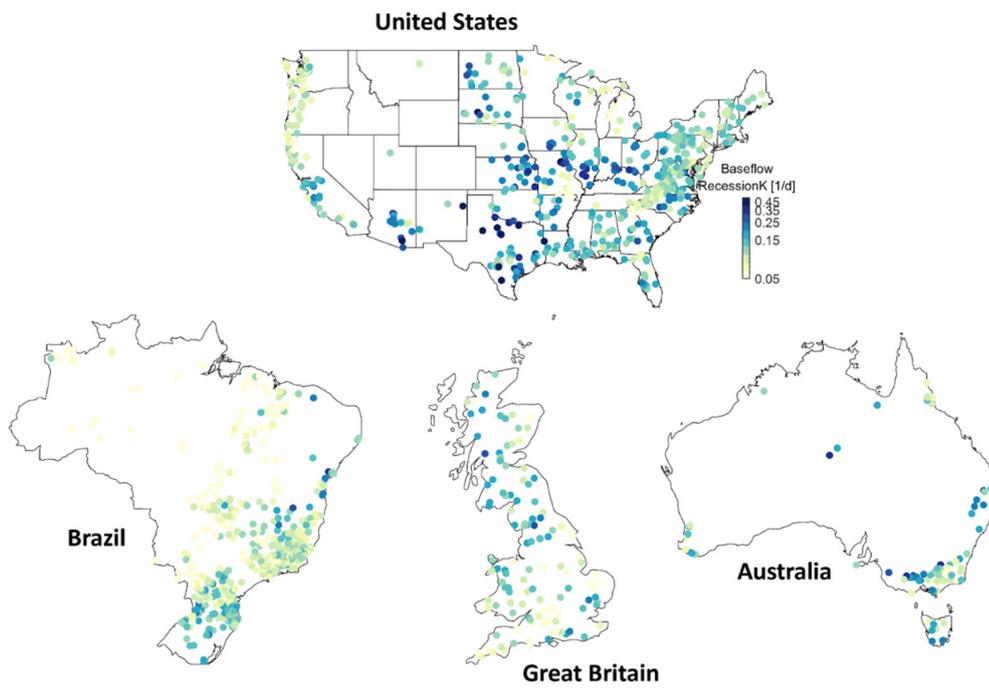


Figure S5. Maps of BaseflowRecessionK. Note that the maps of the countries are not to the same scale.

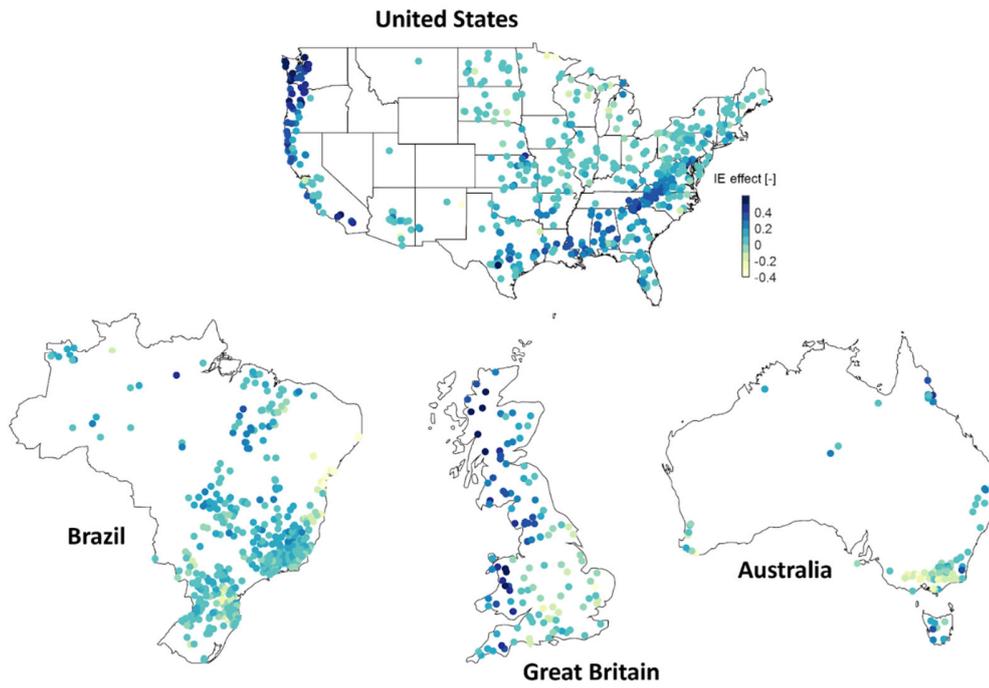


Figure S6. Maps of IE_effect. Note that the maps of the countries are not to the same scale.

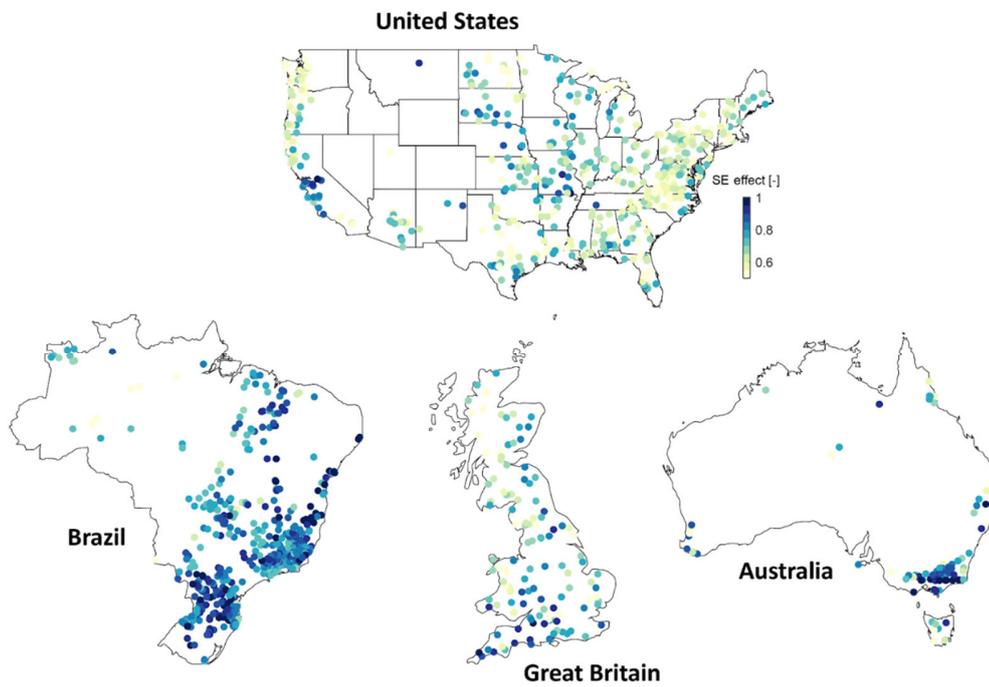


Figure S7. Maps of SE_effect. Note that the maps of the countries are not to the same scale.

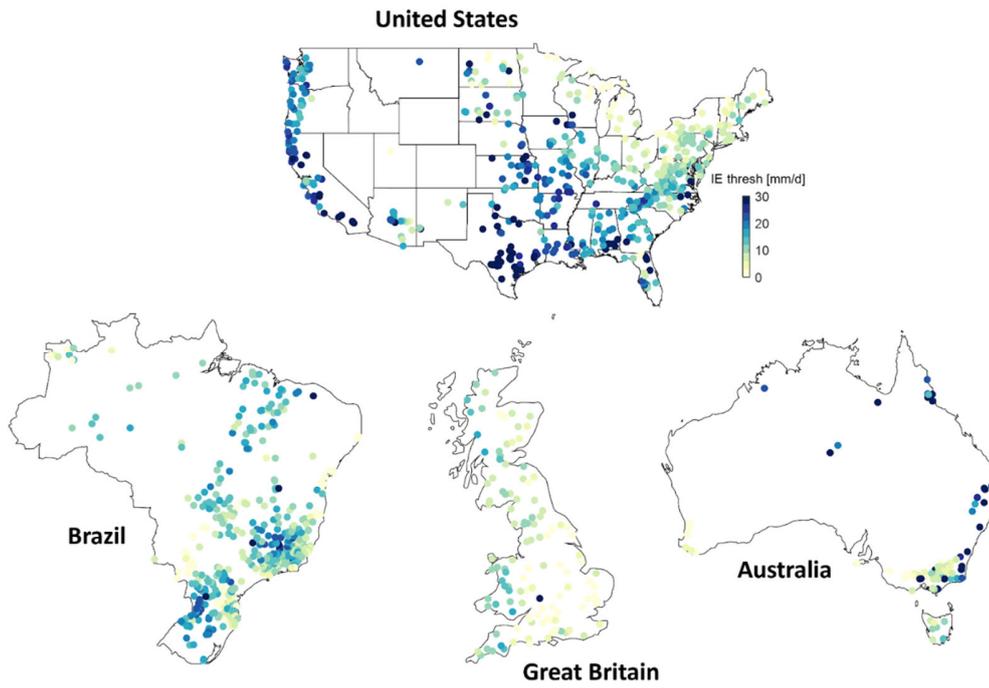


Figure S8. Maps of IE_thresh. Note that the maps of the countries are not to the same scale.

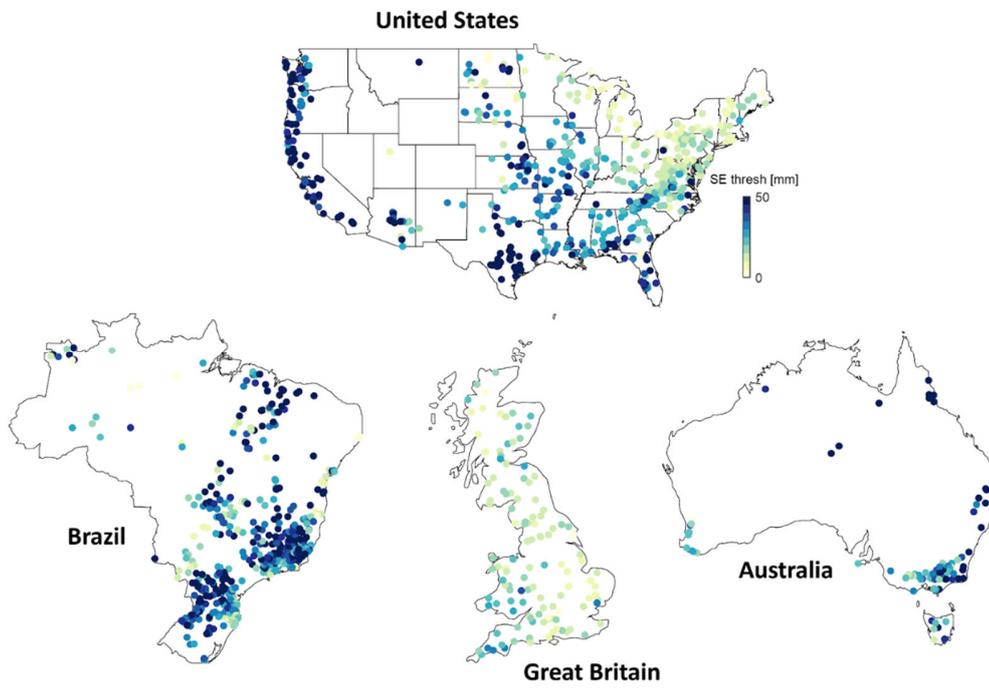


Figure S9. Maps of SE_thresh. Note that the maps of the countries are not to the same scale.

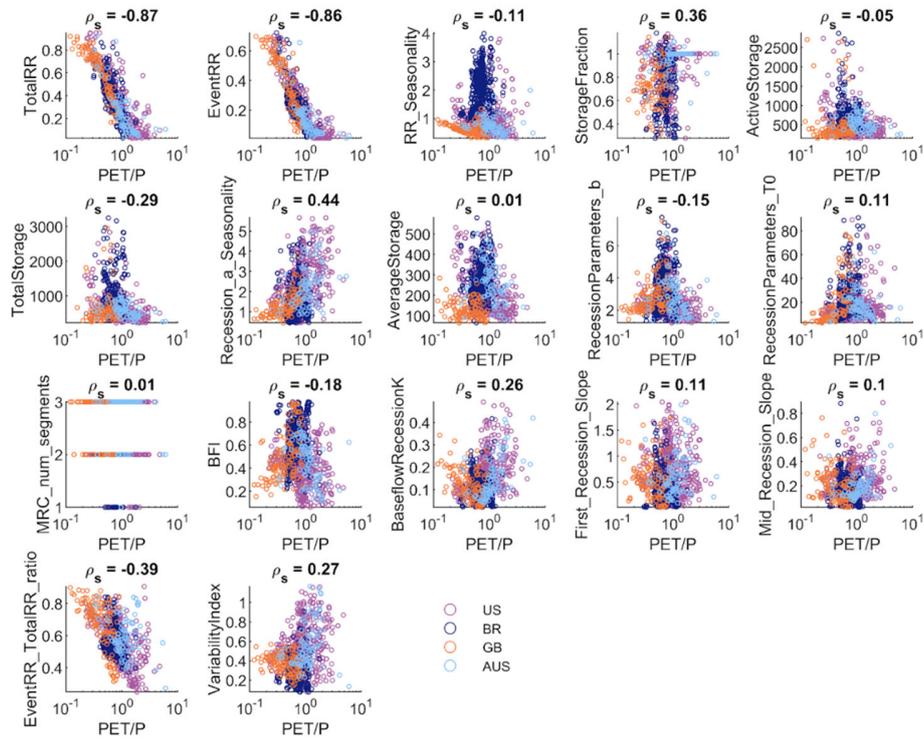


Figure S10. Relationship between groundwater signatures and aridity (PET/P).

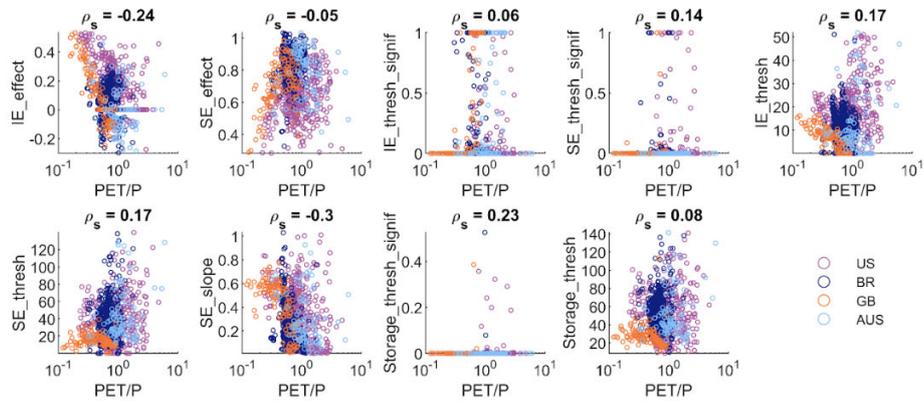


Figure S11. Relationship between overland flow signatures and aridity (PET/P).

Table S1. Signatures that were removed from the analysis.

Signature	Reason
RecessionParameters_a	Units depend on RecessionParameters_b, RecessionParameters_T0 is used instead
Spearman's_rho	Superseded by Recession_a_Seasonality for variations in recessions
min_Qf_perc	Uninformative signature which mostly yields a value of 0

Table S2. Watershed-specific parameters and methods of estimation.

Workflow	Parameter	Description	Method of Estimation
OF	max_recessiondays	Max. length of recession after rainfall, to calculate event volume	Set to 5
GW	recession_length	minimum number of days of decreasing flow required to count as a recession	Set to 2
GW	start_water_year	Month when the water year starts	Set to 10 for US and Great Britain, 4 for Australia, 9 for Brazil
GW	eps	Allowed increase in flow during recession period	Set to median flow * 0.001 in most cases, checked visually that recessions were not being rejected due to diurnal cycles. Higher values (up to median flow * 0.17) needed for intermittent watersheds (Eel and Santa Catalina CZOs).
GW	n_start	Days after flow peak to start recession period	Set to 1

Table S3. 1st, 25th, 50th, 75th, and 99th percentiles of groundwater signatures.

Signature	1st	25th	50th	75th	99th
TotalRR [-]	0.03	0.26	0.37	0.49	0.95
EventRR [-]	0.01	0.14	0.21	0.28	0.72
RR_Seasonality [-]	0.30	0.62	0.96	1.74	4.02
StorageFraction [-]	0.28	0.72	0.91	1.00	1.18
ActiveStorage [mm]	154	347	474	643	2881
TotalStorage [mm]	218	409	568	836	3267
Recession_a_Seasonality [-]	0.44	1.10	1.60	2.46	5.68
AverageStorage [mm]	64	149	207	278	582
RecessionParameters_b [-]	1.20	2.02	2.52	3.46	7.77
RecessionParameters_T0 [d]	2.65	9.54	14.47	22.13	91.31
MRC_num_segments [-]	1.00	2.00	3.00	3.00	3.00
BFI [-]	0.06	0.38	0.53	0.70	0.98
BaseflowRecessionK [1/d]	0.02	0.08	0.11	0.17	0.49
First_Recession_Slope [1/d]	0.01	0.24	0.46	0.77	2.04
Mid_Recession_Slope [1/d]	0.02	0.12	0.17	0.24	0.90
EventRR_TotalRR_ratio [-]	0.25	0.50	0.55	0.64	0.91
VariabilityIndex [-]	0.08	0.25	0.35	0.51	1.17

Table S4. 1st, 25th, 50th, 75th, and 99th percentiles of overland flow signatures.

Signature	1st	25th	50th	75th	99th
IE_effect [-]	-0.30	0.00	0.07	0.18	0.53
SE_effect [-]	0.28	0.62	0.74	0.84	1.03
IE_thresh_signif [-]	0.00	0.00	0.00	0.00	1.00
SE_thresh_signif [-]	0.00	0.00	0.00	0.00	1.00
IE_thresh [mm/d]	0	8	12	18	52
SE_thresh [mm]	0	16	29	45	140
SE_slope [mm/mm]	0.01	0.14	0.26	0.44	1.03
Storage_thresh_signif [-]	0.00	0.00	0.00	0.00	0.53
Storage_thresh [mm]	9	30	44	60	142

Table S5. Rank correlation between groundwater signatures and aridity (PET/P) by country.

Signature	US	GB	AUS	BR	Total
TotalRR [-]	-0.89	-0.90	-0.82	-0.78	-0.87
EventRR [-]	-0.90	-0.92	-0.83	-0.77	-0.86
RR_Seasonality [-]	0.03	-0.48	-0.33	0.48	-0.11
StorageFraction [-]	0.22	0.26	0.27	0.13	0.36
ActiveStorage [mm]	-0.23	0.16	-0.43	0.23	-0.05
TotalStorage [mm]	-0.32	0.11	-0.46	-0.13	-0.29
Recession_a_Seasonality [-]	0.36	0.67	0.41	0.25	0.44
AverageStorage [mm]	-0.38	-0.17	-0.48	0.38	0.01
RecessionParameters_b [-]	-0.24	0.47	-0.45	0.22	-0.15
RecessionParameters_T0 [d]	0.11	0.77	-0.28	0.36	0.11
MRC_num_segments [-]	-0.06	-0.34	0.11	-0.09	0.01
BFI [-]	-0.25	0.54	-0.56	0.29	-0.18
BaseflowRecessionK [1/d]	0.38	-0.40	0.62	-0.18	0.26
First_Recession_Slope [1/d]	-0.05	-0.27	0.29	-0.22	0.11
Mid_Recession_Slope [1/d]	0.16	-0.30	0.44	-0.29	0.10
EventRR_TotalRR_ratio [-]	-0.59	-0.65	-0.06	-0.20	-0.39
VariabilityIndex [-]	0.26	-0.27	0.39	-0.18	0.27

Table S6. Rank correlation between overland flow signatures and aridity (PET/P) by country.

Signature	US	GB	AUS	BR	Total
IE_effect [-]	-0.34	-0.92	-0.18	0.03	-0.24
SE_effect [-]	0.25	0.40	-0.25	0.04	-0.05
IE_thresh_signif [-]	0.24	0.87	0.12	0.17	0.06
SE_thresh_signif [-]	0.06	0.57	0.35	0.15	0.14
IE_thresh [mm/d]	0.24	-0.79	-0.02	-0.12	0.17
SE_thresh [mm]	0.26	-0.26	0.05	0.11	0.17
SE_slope [mm/mm]	-0.32	-0.73	-0.29	-0.54	-0.30
Storage_thresh_signif [-]	0.19	0.61	0.50	0.12	0.23
Storage_thresh [mm]	0.17	-0.51	-0.01	0.09	0.08